Cb-LIKE - Cumulonimbus Likelihood: Thunderstorm forecasting with fuzzy logic

Martin Köhler*, Arnold Tafferner* and Thomas Gerz*

Martin Köhler:
German Aerospace Center (DLR), Institute for Atmospheric Physics
Münchner Str. 20, 82234 Oberpfaffenhofen, Germany
phone: +49 8153 28-3545
e-mail: martin.koehler@dlr.de

Arnold Tafferner:
German Aerospace Center (DLR), Institute for Atmospheric Physics
Münchner Str. 20, 82234 Oberpfaffenhofen, Germany
Phone: +49 8153 28-1218
e-mail: arnold.tafferner@dlr.de

Thomas Gerz:
German Aerospace Center (DLR), Institute for Atmospheric Physics
Münchner Str. 20, 82234 Oberpfaffenhofen, Germany
Phone: +49 8153 28-1333
e-mail: thomas.gerz@dlr.de

Abstract

Aviation is heavily affected by thunderstorms. Approaching storm cells with accompanying effects like heavy rain, hail or downdrafts cause delays and flight cancellations. As a negative result the airlines and airport operators have to bear high additional costs. Moreover, also flight's safety and passengers' comfort are reduced. A reliable thunderstorm forecast up to several hours ahead saves time for decision makers as airport authority, air traffic control, airline operation centre, and the crew in the cockpit for appropriate and harmonised reaction and for initiating adequate counteractions to mitigate the consequences of a thunderstorm. The algorithm Cb-LIKE (**C**umulonim**b**us-**Like**lihood) has been developed to provide such forecasts. Cb-LIKE is an automated system which designates areas with possible thunderstorm development by using output of the COSMO-DE numerical weather prediction model operated by the German Meteorological Service (DWD). The algorithm includes a newly developed "Best-Member-Selection" which allows the automatic selection of that member of a COSMO-DE ensemble that matches best the current weather situation. An innovative fuzzy logic system combines selected model data and calculates a thunderstorm indicator for each grid point of the model domain for the following six hours in one hour intervals. Comparing thunderstorm observations by radar with Cb-LIKE forecasts in the summer period of 2012 verifies the algorithm and demonstrates the system's performance. Moreover, the verification results allow to transform the Cb-LIKE indicator field into a field of thunderstorm probability.

Keywords

Nowcasting, forecasting, thunderstorms, fuzzy logic, numerical weather prediction, best member, ensemble, verification, seamless prediction

1 Introduction

Thunderstorms are one of the most exciting meteorological phenomena. They can be mainly observed on hot summer days when the atmosphere features perfect conditions for the development of enormous and deep convective cumulonimbus clouds. But not only meteorologists, almost all people are attracted to thunder, hail, lightnings, heavy rain, downdrafts or, in rare cases, tornados. Especially, the air transport system can be massively affected by storm cells. From 40 % up to 50 % of delays in Europe are caused by adverse weather events (EUROCONTROL, 2011) whereof thunderstorms have a large share. According to EUROCONTROL (2011), weather phenomena which bring about the biggest impact on delays within the European aviation are storm cells (30.9 %), snow and ice (22.1 %), low visibility and cloud ceiling (21.5 %) and also wind (14.5 %). In the USA thunderstorms are responsible for even up to 90 % of delays in aviation during the summer months (LEIGHTON, 2006). Delays and also flight cancellations cause high additional costs for airlines and airport operators. Besides there is also the problem of reduced comfort for aviation customers.

Due to these negative impacts there is certainly a demand for thunderstorm prognoses of high quality. In the last decades scientists developed and applied many different methods and techniques which can be grouped in two categories: deterministic/short-term prognoses (0-1 h: nowcasting) on the one hand and probabilistic/long-term prognoses (1 h - several days: forecasting) on the other. The latter category includes approved methods like the use of sounding data which enables an estimation of the atmospheric conditions with regard to the development of thunderstorms (e.g. MUELLER et al., 1993; MANZATO, 2005). Another technique is the usage of a "Model Output Statistics" (MOS) system for the incorporation of "Numerical Weather Prediction" (NWP) model forecasts in traditional statistical methods (WILKS, 2006). Application examples of this approach in long-term storm forecasting can be found in TREPTE (2011) or SCHMEITS et al. (2005). A new technique is shown in KOBER et al. (2012) where a probabilistic

nowcasting method is combined with a high-resolution NWP model ensemble, facilitating probabilistic forecasts of convective events up to eight hours. Besides there are also many approved techniques applied in nowcasting of storms. A well-known tool is the ”Auto Nowcast System” (ANC) of the ”National Center for Atmospheric Research” (NCAR; MUELLER et al., 2003). This algorithm combines several data sources like radar/satellite data, soundings and more within a fuzzy logic system. Another possible approach is the usage of real-time lightning data which is presented in BETZ et al. (2008).

In general, for several decades the most applied technique for deterministic nowcasting of thunderstorms has been the extrapolation of radar/satellite data in time. Various algorithms like TITAN (DIXON and WIENER, 1993), WDDS (EILTS and COAUTHORS, 1996) or SWIRLS (WONG et al., 2006) are based on this approach. Two of the more recent tools in this field are Rad-TRAM (”Radar Tracking And Monitoring”; KOBER and TAFFERNER, 2009) and Cb-TRAM (”Cumulonimbus Tracking And Monitoring”; ZINNER et al., 2008) which provide high-quality deterministic nowcasting of storm cells based on radar and satellite data.

However, nowcasting of thunderstorms based on observations are limited to forecast horizons of one hour at most. On the other hand, decision makers in the aviation business as airport authority, air traffic control, airline operation centre, and the crew in the cockpit all claim the necessity of storm forecasts for some hours ahead for appropriate and harmonised reaction and for initiating adequate counteractions to mitigate the economic consequences of a thunderstorm and to preserve safety of the flight and comfort of the passengers. Therefore, an algorithm is required which provides high-quality storm forecasts up to several hours and connecting seamlessly to a nowcasting tool. To this end, Cb-LIKE has been developed at the DLR (Deutsches Zentrum für Luft- und Raumfahrt) Institute of Atmospheric Physics.

A fuzzy logic approach is the heart of the new algorithm in Cb-LIKE. This technique combines particular NWP model output parameters and delivers the likelihood of thunderstorms

in terms of an indicator. In general, a fuzzy logic system allows a fast, realistic, problem related, and significant modelling of complex systems with nonlinear behaviour. Linguistic expressions can describe the system which makes it clearly easier in comparison to mathematical description methods. It allows to solve problems which arise from vague, ambiguous, incomplete or imprecise information (MURTHA, 1995). This is in particular valid for thunderstorm prognoses on the basis of model runs where unresolved features like soil moisture or convection have to be parameterized. Another issue are boundary and initial conditions of the numerical weather models. The use and the benefit of the application of a fuzzy logic system for meteorology are illustrated for example in the publications of MUELLER et al. (2003), HANSEN (2007) or KEIS (2015). The way of applying fuzzy logic solely to model data is an entirely new approach to forecast thunderstorms probabilistically several hours ahead. Another innovative part of the Cb-LIKE algorithm is a direct conjunction to an already existing nowcasting approach. This procedure, named ”Best-Member-Selection”, allows an automatic selection of the best initial data basis to calculate the thunderstorm forecasts.

The data sources used in the Cb-LIKE algorithm are introduced briefly in section two. Section three contains an overview of the newly designed fuzzy logic system. Four case examples which give insight in the forecasting quality of the new algorithm are shown in chapter four. The ”Best-Member-Selection” is described in detail in section five. Chapter six includes some results of a thorough verification of Cb-LIKE for the full summer period of 2012.

2 Data sources

We use the output of the COSMO-DE model of the German Meteorological Service (DWD) for the calculation of the thunderstorm forecasts within the Cb-LIKE algorithm. Therefore this model can be seen as the main data source. It is a nonhydrostatic numerical weather model with a resolution in the meso-γ scale (Δx $\sim$ 2.8 km) (SCHÄTTLER et al., 2013). Within Middle Europe it possesses one of the highest resolutions among all available operational models. It provides forecasts up to 21 hours with a 3 hour update rate between 0000 and 2100 UTC and the domain covers Germany and parts of the neighbouring countries (figure 1). Its grid comprises 421x461 points and 50 vertical height levels. In contrary to its precursors (e.g. COSMO-EU; SCHULZ and SCHÄTTLER, 2010) the COSMO-DE model features a full resolution of large convective phenomena (no parameterization of deep convection) as a result of its high grid resolution (BALDAUF et al., 2011). As a consequence it is a logical choice for the application in operational thunderstorm forecasting in Middle Europe.

The Rad-TRAM algorithm serves as second data source within Cb-LIKE and is used for the Best-Member-Selection and also for the verification of the Cb-LIKE prognoses. Rad-TRAM was developed at the DLR Institute for Atmospheric Physics and stand for "**Rad**ar **T**racking **a**nd **M**onitoring" and allows a reliable detection, tracking and nowcasting (0-1 h) of heavy precipitation cells (KOBER and TAFFERNER, 2009) using the European radar composite issued by the DWD. This consists of radar reflectivities given in six dBZ classes with a horizontal resolution of 2 km x 2 km and encompasses an area of 1800 km x 1800 km (WEIGL et al., 2005) as illustrated in figure 2. A heavy precipitation cell must consist of at least 21 contiguous pixels and features therefore a minimum size of 81 km^2 (ZINNER et al., 2008). The reflectivity values of the radar composite are composed of measurements from 3-dimensional scans of various radars across Central Europe. Due to the fact that the Rad-TRAM domain covers nearly the complete COSMO-DE model area and renders the current storm situation, its application within the "Best-

Member-Selection" and the verification of the Cb-LIKE forecasts is a sensible decision.

3 Fuzzy logic system

In general, a fuzzy logic system consists of three different main steps (figure 3): fuzzification, inference and defuzzification. In the fuzzification step membership functions assign the incoming crisp variables (that is a parameter with a value, e.g. CAPE = 500 J/kg) to membership grades ranging from 0 to 1. These functions represent the so-called fuzzy input-sets of the fuzzy logic system and are always named by linguistic variables (e.g. ”low”, ”moderate”, ”high”). Hence, the fuzzy logic approach turns away from binary logic where only 0 and 1 exist and allows a statement over a parameter value to be partially true or false. Referring to Cb-LIKE, the used model parameters stand no longer solely pro or contra to thunderstorm development. They can now be rated as for example ”partly pro” or ”rather contra”. Thereby problems associated with traditional binary logic can be avoided and the solutions are more oriented to human reasoning. In the fuzzy inference step an ”if...then” decision rule base combines the membership grades of various input parameters (as defined by the fuzzy input sets) yielding fuzzy output sets. The decision rule base has to be developed first, thereby using facts and experiences from the ”knowledge base” (figure 3). Thirdly, the defuzzification step calculates the final output from the fuzzy output sets. This represents again a crisp parameter (hence, a parameter with a value), in case of Cb-LIKE a so-called ”thunderstorm indicator” with a value between 0 and 100. The fuzzy logic system adapted to the requirements of Cb-LIKE is presented in the following, mostly refering to JANTZEN (1998) or ROSS (2010).

We apply four COSMO-DE model parameters in the fuzzy logic system which are CAPE (**C**onvective **A**vailable **P**otential **E**nergy), Omega on the 500 hPa level (vertical wind), the radar reflectivity (dBZ) and the cloud top temperature (IR 10.8, derived from satellite data). In literature these four are well-known and often used parameters with regard to thunderstorm detection as well as forecasting and include three different parameters in meteorology. CAPE gives information about the state of the atmosphere whereas the vertical wind Omega is used as an

indicator for the initiation of thunderstorms. The radar reflectivity and cloud top temperature are often used for the detection of already developed thunderstorm cells. Hence these four parameters are predestinated for the application in Cb-LIKE.

3.1 Fuzzy input sets - Fuzzification

The fuzzy input sets developed for the fuzzification step are illustrated in figure 4 (a) - (d). The respective value range of each parameter is depicted on the abscissa whereas the ordinate comprises the membership grade between 0 and 1 in each figure. Three different membership functions named by the linguistic variables ”low”, ”moderate” and ”high” are used for each parameter. An outline of the value ranges comprised by all membership functions can be found in table 1. Among many possible forms (see for example KLIR and YUAN, 1995) we apply for all sets standard trapezoidal shapes which are often used in meteorological fuzzy logic systems like for example MURTHA (1995) or KEIS (2015). The transition areas between the individual membership functions, which enable the fuzzification of the crisp input parameters, are kept symmetrical. For example a value of CAPE = 500 J/kg would be assigned to the sets low and moderate with a respective membership grade of 0.5. The fact that the assignment of the input values to the fuzzy sets is not limited by the value range of the crisp parameter is a further advantage of the application of a fuzzy logic system because an input value lying beyond the value range will automatically be assigned to the lowest or highest set. For instance, the fuzzy input sets which are defined for the vertical movement Omega do not comprise positive values and therefore do not cover downdrafts at a first glance. But if a grid point possesses a positive Omega it will automatically be assigned to the fuzzy set ”low”. This set indicates rather no storm development and therefore downdrafts are taken into account within the Cb-LIKE algorithm.

3.2 Fuzzy logic inference - Rule base

The connection between the fuzzy input and output sets (see chapter 3.3) is provided by the fuzzy inference step. It presumes the construction of a so-called rule base consisting of several "if...then" decision rules. The "if" portion of a rule refers to the membership grade in one of the fuzzy input sets whereas the "then" portion refers to the "consequences" in the associated output set. For example, the strongest rule in Cb-LIKE concerning thunderstorm development consists of the following "if...then" combination:

- If (CAPE$_{high}$ & Omega$_{high}$ & Radar$_{high}$ & CloudTop$_{low}$) Then Indicator$_{veryhigh}$

These rules, which have to be defined for each fuzzy system, of course should reflect physical relations but they also depend strongly on the expert knowledge of the developer. The total amount of rules is the product of the number of fuzzy input sets describing the fuzzy system. In case of Cb-LIKE, three fuzzy input sets for each of the four parameters lead to the total amount of $3^4 = 81$ rules. For a structured conjunction of the rules with the output sets, we applied the approach of a weighted average in Cb-LIKE. For this purpose a "virtual" score is first assigned to each fuzzy input set, +1 for sets which indicate thunderstorm development, 0 for neutral and -1 for fuzzy sets indicating no storm development, see table 2. The next step is the calculation of the weighted average m for each rule by

$$m = \frac{\sum_{i=1}^{n} x_i g_i}{\sum_{i=1}^{n} g_i}, \tag{3.1}$$

where x_i stands for the virtual score of each fuzzy input set, g_i denotes the associated weighting and i moves over all sets combined within a rule, hence $n = 4$ in our case here. Within this setting of Cb-LIKE all fuzzy input sets are equally weighted ($g_i = 1$). Therefore, the weighted average m can attain values from -1 up to +1 for any of the 81 "if...then" decision rules. Its general conjunction to the fuzzy output sets is illustrated in table 3. For example the strongest

rule indicating storm development as shown before attains a weighted average m of +1.0 and is linked to the output set "very high". An overview of the symmetrical distribution of all 81 rules to the five output sets dependent on m can be found in table 4.

3.3 Fuzzy output sets - Defuzzification

The fuzzy output sets are illustrated in figure 5. The ordinate comprises again the membership grade from 0 up to 1 whereas the abscissa extends over a value range from 0 up to 100 which represents the "thunderstorm indicator" and is covered by five membership functions. Two of them feature the shape of a triangle ("low" & "high") the other three possess a trapezoidal form ("very low", "moderate", "very high"). The transition areas between the sets are again kept symmetrical. Note that the minimum and maximum values of the indicator are not 0 and 100 but $11.\bar{6}$ and $88.\bar{3}$, respectively, resulting from the centre-of-gravity rule applied to the half trapezoidal shape of the two extreme output sets "very low" and "very high". For the calculation of the final indicator the membership grade of each output set has to be determined by the weighted and assigned decision rules as described before. In Cb-LIKE the Root-Sum-Square (RSS) (e.g. UMOH et al., 2010) approach is applied for this step,

$$\mu_j = \sqrt{\sum_{i=1}^{n} R_i^2}. \tag{3.2}$$

R_i is the strength of each decision rule. The "strength" depends on the weakest "if" portion which is the degree of membership of each input parameter to the corresponding fuzzy input set(s). Therefore, R_i represents the strengths from 0 up to 1 of the n rules which are linked to the same fuzzy output set (table 3). Hence this approach enables the best weighted membership grade μ_j between 0 and 1 of each fuzzy output set j under the influence of all affected ("firing") rules.

For the final calculation of the thunderstorm indicator we used the "Weighted Average Method" (e.g. ROSS, 2010),

$$z = \frac{\sum_{j=1}^{m} \mu_j * w_j}{\sum_{j=1}^{m} \mu_j}, \tag{3.3}$$

where w_j represents the centre of gravity of the output membership function j. For Cb-LIKE, 5 output sets are defined, hence $m = 5$. z denotes the finally calculated weighted centre of gravity of all weighted input membership functions. Its x-value finally represents the wanted crisp output, in case of Cb-LIKE the thunderstorm indicator which varies between $11.\bar{6}6$ und $88.\bar{3}3$.

4 Case studies

The following section presents the functional principle of Cb-LIKE. At first the impact of the four different input parameter values on the calculation of the thunderstorm indicator is illustrated in three examples. Afterwards a real case including up to six hours of thunderstorm forecasting is presented in order to demonstrate the performance of Cb-LIKE.

4.1 Example case: Impact of different input parameter values

In order to show the functional principle of Cb-LIKE in connection with input parameters indicating rather no thunderstorm development, they are set on the values CAPE = 450 J/kg, Omega = -45 hPa/h, Radar reflectivity = 23 dBZ, and Cloud top temperature = 260 K, respectively.
From figures 4 (a) - (d), we see that these input parameters are consequently assigned to the fuzzy input sets "low" and "moderate" for CAPE, Omega and radar reflectivity and to "moderate" and "high" for cloud top temperature. The Cb-LIKE algorithm calculates a resulting "low" thunderstorm indicator (value 29.5) as illustrated in figure 6 (a), where the five fuzzy output sets attain the membership grades "very low" (0.42), "low" (0.96), "moderate"(0.27), "high" (0), and "very high" (0) according to the rule base.

The second example demonstrates the performance of Cb-LIKE when the input parameters CAPE = 950 J/kg, Omega = -98 hPa/h, Radar reflectivity = 41 dBZ, and Cloud top temperature = 228 K indicate that a thunderstorm is likely to develop. The assigned fuzzy input sets read ”moderate/high” for CAPE, Omega and radar reflectivity and ”low/moderate” for the cloud top temperature, resulting in a rather high thunderstorm indicator of 67.1 as illustrated in figure 6 (b). According to the rule base the five fuzzy output sets attain the membership grades ”very low” (0), ”low” (0), ”moderate”(0.48), ”high” (0.85), and ”very high” (0.32). Due to the highest membership grade of the output set ”high” and a higher grade of the set ”moderate” in comparison to ”very high”, the averaged value over all fuzzy output sets attains a x-value of 67.1.

The third example demonstrates the behaviour of Cb-LIKE when some input parameters indicate a storm development and others don’t, as outlined in the values of CAPE = 450 J/kg, Omega = -98 hPa/h, Radar reflectivity = 41 dBZ, and Cloud top temperature = 260 K. An inconsistent value distribution like this (two of four parameters indicate a possible storm development) is not reasonable regarding the real atmosphere but can occur as a consequence of the inadequate model physics.

The assigned fuzzy input sets then read ”low/moderate” for CAPE, ”moderate/high” for Omega and the radar reflectivity and ”high/moderate” for the cloud top temperature. Cb-LIKE calculates a rather neutral thunderstorm indicator of 46.1 as illustrated in figure 6 (c) with the membership grades of the five fuzzy output sets of ”very low” (0), ”low” (0.87), ”moderate”(0.64), ”high” (0.48), and ”very high” (0).

4.2 Example case: 22 June 2011

Now, we demonstrate the performance of Cb-LIKE to forecast thunderstorms up to six hours ahead of time. In the afternoon of 22 June 2011, a distinct upper level trough caused high thunderstorm activity over Central Europe. The axis of the trough extended from the British

Islands across France such that Germany and parts of the neighbouring countries lay on its forefront which is generally characterized by a destabilized atmosphere. We used the 1200 UTC COSMO-DE model-run for Cb-LIKE to forecast storm indicators between 1300 and 1800 UTC in one hour intervals. In figure 7 the forecasted thunderstorm indicators are plotted as coloured surfaces together with heavy precipitation cells as observed by Rad-TRAM (threshold 37 dBZ) at the respective time to verify the Cb-LIKE forecasts.

Figure 7 (a) displays major Rad-TRAM objects west of the French-Swiss border and over central and northern Germany at 1300 UTC. The associated one hour forecast of Cb-LIKE shows a very good accordance of coloured surfaces including high indicator values up to 80 to the observations west of the French-Swiss border. In central and northern Germany the highest indicator values are slightly displaced to the east in relation to the observed Rad-TRAM objects. At 1400 and 1500 UTC (figures 7 (b) and (c)) the thunderstorm activity over Central Europe shifts towards east. Over central and northern Germany the Rad-TRAM objects now overlap very well with high indicator surfaces of Cb-LIKE. A further area of heavy precipitation is observed south-east of Munich and well forecast by Cb-LIKE with indicator values of 60/70 for 14 UTC and 80 for 15 UTC, respectively. Also the thunderstorm activity at the French-Swiss border is still satisfactorily covered by high indicator values. At 1600 UTC (figure 7 (d)) the thunderstorm activity has further moved towards north-east. The four hour forecast of Cb-LIKE still shows high indicator values in areas where heavy rain is observed, especially over north-eastern Germany and on half way between Munich and Vienna. The Cb-LIKE forecasts for northern Switzerland and south-western Germany are also quite good. Even the five and six hours forecasts of Cb-LIKE (for 1700 and 1800 UTC, see figures 7 (e) and (f)) are quite satisfying in general. Nevertheless, in some areas, as for example over northern Switzerland and South-West Germany, the highest indicator values now show a noticeable spatial displacement in comparison to the observed storm activity.

In summary the Cb-LIKE algorithm shows a satisfying forecasting performance for the chosen example case. The four input parameters used to compute the indicator are shown separately in figures 8 (a) - (d) together with the Rad-TRAM observations. The parameters base on the 1200 UTC COSMO-DE model run and are valid for 1600 UTC. The respective Cb-LIKE forecast is displayed on figure 7 (d). Each input parameter denotes very large areas in which a development of thunderstorms is likely. These areas are much bigger than the corresponding Cb-LIKE prognoses. Additionally, all four input parameters tend to overforecast areas with possible thunderstorm development especially in the eastern (CAPE) and north/north-western parts (OMEGA, Cloud top temperature, Radar reflectivity) of the model domain. Hence, only the fuzzy combination of the parameters features a better forecast in terms of less overforecasting and smaller areas of thunderstorm developments.

5 Best-Member-Selection

TAFFERNER et al. (2002) proposed the idea of a ”Best-Member-Selection” which was later applied to satellite image matching by TAFFERNER et al. (2008). A variant of this technique is developed here for Cb-LIKE in order to provide the best possible COSMO-DE model output for the fuzzy logic system. The ”Best-Member-Selection” automatically selects that model run out of a time-lagged ensemble which matches best the current observed weather situation. Due to the fact that the latest (youngest) model run does not always provide the best depiction of the current thunderstorm situation and therefore does not offer the best starting position for calculating thunderstorm prognoses, the implementation of the ”Best-Member-Selection” is a sensible step. Since the maximum forecast lead time of COSMO-DE is 21 hours with a new start each 3 hours, the time-lagged ensemble is limited to the last five model runs to ensure thunderstorm predictions up to six hours. If for example the Cb-LIKE forecasts are calculated at 1400 UTC time, the model runs issued at 1200, 0900, 0600, 0300 and 0000 UTC will be available and used

to search for the best ensemble member compared to the current thunderstorm situation. The thunderstorm situation is thereby represented by heavy precipitation cells as they are detected by Rad-TRAM out of the European radar composite with a minimum threshold of 37 dBZ. For a sensible comparison between the model runs and Rad-TRAM data, synthetic heavy precipitation cells are computed from the COSMO-DE radar-field output using the same threshold. Finally the synthetic precipitation cells of a model run which match best the cells observed by Rad-TRAM define the best member. Generally speaking, the "Best-Member-Selection" serves as an interface allowing to combine nowcast and forecast techniques to provide seamless predictions over many time horizons.

Altogether the "Best-Member-Selection" implemented in Cb-LIKE offers four different modes to select the best ensemble member. They read as follows:

Object comparison. The first mode provides an object comparison of the COSMO-DE synthetic radar data of the time-lagged ensemble with Rad-TRAM objects. For each model run the number of synthetic radar objects which overlap with Rad-TRAM objects at the current point of time is determined. A single overlapping grid point is sufficient to count the synthetic radar object as a match. That model run of the time-lagged ensemble with the most matching objects will automatically be chosen for the application in the fuzzy logic system. If two or more model runs possess the same number of matches, the latest run will be selected.

Grid point comparison. In this case, the selection criterion for the best fitting model run is the highest number of matching grid points between forecast (synthetic radar cell grid points) and observation (grid points occupied by Rad-TRAM objects). This comparison is very strict, as it counts matching grid points only.

Object ratio. The third mode calculates the object ratio $O_r = O_o/O_a$ to select the best fitting COSMO-DE model run. O_o represents the total number of synthetic radar objects of a model run which overlap at least at one grid point with Rad-TRAM objects and O_a is the number of all

synthetic radar objects of the run. That member of the ensemble is chosen as best which possesses the highest object ratio O_r. In contrast to the "classical" object comparison described above, the O_r method is more rigorous with wrongly forecast synthetic radar objects: The more synthetic radar objects exist without any overlap to Rad-TRAM objects the smaller O_r will become and the respective ensemble member will be sorted out. Or, inversely, when only a few synthetic radar objects exist in the domain and some of them overlap with observations, O_r will become quite large.

Grid point ratio. The fourth mode utilizes the grid point ratio $G_r = G_o/G_a$ to filter out the best COSMO-DE model run, where G_o is the number of grid points of matching synthetic radar objects of a model run and G_a signifies all grid points occupied by synthetic radar objects. Similar to O_r, also G_r is more strict compared to the "classical" grid point method. Since it counts grid points instead of objects this method is the most rigorous one of all the four introduced methods.

5.1 Example case: 15 August 2012 - 1500 UTC

The following example case illustrates the advantage of the application of the "Best-Member-Selection" within the Cb-LIKE algorithm. Figures 9 (a) to (e) show the thunderstorm situation in Upper Bavaria at 1400 UTC on 15 August 2012. One blue polygon represents a small heavy precipitation cell observed by Rad-TRAM in the southwest of the German Aerospace Centre in Oberpfaffenhofen (marked with "DLR"). The coloured surfaces show the COSMO-DE synthetic radar reflectivity fields of a specific model run of the time-lagged ensemble, respectively. In figure 9 (a) we display the radar reflectivity of the most recent model run from 1200 UTC. There is almost no match of reflectivity values higher than 19 dBZ and the observation. Therefore, the two-hour prognosis gives a rather weak indication of the observed thunderstorm situation at 1400 UTC. The radar reflectivity field from the 0900 UTC run shows better results (figure 9 (b)) as a larger area of reflectivity values higher than 19 and even higher than 28 dBZ matches with the

southern part of the observed heavy precipitation cell. In contrast the synthetic radar reflectivity fields of the model runs from 0600 and 0300 UTC (figure 9 (c) - (d)) show only a small match of dBZ values between 7 and 19 with the observed Rad-TRAM cell. Finally, the oldest model run from 0000 UTC yields no reflectivity values higher than 7 dBZ in accordance with the observed Rad-TRAM cell. In conclusion, the COSMO-DE model run from 0900 UTC represents best the current weather situation at 1400 UTC in comparison to all other model runs of the time-lagged ensemble.

As a consequence, the calculation of the Cb-LIKE indicators starting at 1400 UTC and applying the ”Best-Member-Selection” (using the Grid-Point-Ratio or Grid-Point Comparison) should lead to a better forecasting performance one hour later at 1500 UTC. Figures 10 (a) - (e) display the thunderstorm situation in Upper Bavaria at 1500 UTC on 15 August 2012. Two blue polygons represent a small and a large heavy precipitation cell observed by Rad-TRAM in the southwest of the German Aerospace Centre in Oberpfaffenhofen (marked with ”DLR”). The coloured surfaces show the Cb-LIKE indicator fields using data from the last five COSMO-DE model runs of the time-lagged ensemble. Indeed, the Cb-LIKE prognosis based on the 0900 UTC run (figure 10 (b)) shows the best performance as two areas with indicator values higher than 50 are calculated whereof one matches with the southern part of the observed area. The Cb-LIKE forecast based on the model run from 0300 UTC shows a slight overlap with the large Rad-TRAM cell with indicator values between 40 and 60. The prognoses based on the 1200, 0600 and 0000 UTC runs, respectively, show overlaps with only low indicator values between 30 and 40. Notably, the Cb-LIKE forecast based on the latest model run features the worst performance. To conclude, the ”Best-Member-Selection” is a useful tool, resulting in a better Cb-LIKE prognoses, since the latest model run does not necessarily provide the best data basis for a thunderstorm forecast.

When the domain, for which the Cb forecast is applied, is large, like for example the area

of Germany, it is often not clear which model run matches best the current weather situation. Thunderstorms may occur at the same time in different regions which are far apart from each other. Hence, different model runs can simultaneously match well the thunderstorm situation in one of the regions and fail to predict the Cb activity in other regions. In such a case it is not well defined which thunderstorm activity is more important to be forecasted for a possible user and which individual model run is the best to be selected automatically. Therefore, the "Best-Member-Selection" should be applied when the focus lies on small domains (e.g. an airport environment or a city area) in which thunderstorms develop or don't develop for the same meteorological reasons.

6 Verification of Cb-LIKE

Verification is an indispensable part of meteorological research and operational forecasting activities. If the methodology is properly designed, verification results can effectively meet the needs of many diverse groups, including modellers, forecasters and users of forecast information (Casati et al., 2008). We now describe the performance of Cb-LIKE in the framework of a verification exercise in the summer period of 2012 for the area of Germany. For this exercise the "Best-Member-Selection" is not used.

We compare Cb-LIKE objects with heavy precipitation cells as detected by Rad-TRAM and forecasted by the COSMO-DE synthetic radar field which can also be used as thunderstorm forecast. By comparing both prognosis fields, we evaluate if the new algorithm represents an improvement in terms of thunderstorm forecasting on the basis of COSMO-DE model data.

Before giving an outline of the "Neighbourhood Verification", the verification scores which are used are introduced briefly. The definite forecast objects from Cb-LIKE and the synthetic radar data represent binary (yes/no or 0/1) thunderstorm prognoses. These are also called "dichotomous forecasts", see for example DOSWELL III et al. (1990). So-called 2x2 contingency

tables are traditionally applied for the verification of these forecasts (see figure 11). According to e.g. WILKS (2006), these kind of tables include four possible forecast/observation pairs which are described as

- ”Hit” - The event was successfully forecasted to occur;
- ”False alarm” - The event was forecasted to occur but did not occur;
- ”Miss” - The event occurs but was not forecasted;
- ”Correct negative” - The event did not occur after a forecast that it would not occur.

The usage of these four forecast/observation pairs is a useful way to show the errors caused by a forecasting algorithm. A perfect prognosis would only produce ”hits” and ”correct negatives”.

On the basis of 2x2 contingency tables many different verification scores can be calculated which describe different aspects of forecasting quality. Here we introduce four of them.

The first one is the ”False Alarm Ratio” (FAR). This quantity

$$\textbf{FAR} = \frac{false\ alarms}{hits + false\ alarms} \tag{6.1}$$

measures the fraction of forecasted events which do not occur. Its value range comprises 0 up to 1, a perfect forecast would attain a FAR = 0.

Additionally the ”Probability of Detection” (POD) is calculated by

$$\textbf{POD} = \frac{hits}{hits + misses}. \tag{6.2}$$

This parameter represents the portion of observed events which are correctly forecasted. It also covers a value range from 0 up to 1 and a perfect forecast would attain a POD = 1.

The third verification score is the BIAS

$$\textbf{BIAS} = \frac{hits + false\ alarms}{hits + misses}, \tag{6.3}$$

which measures the ratio of forecasted to observed events. It encompasses a value range from 0 up to ∞ whereas its perfect value is 1. The BIAS indicates if a forecasting algorithm calculates too few (BIAS $<$1) or too many (BIAS $>$1) objects in comparison to the number of observed objects.

The forth verification score is the ”Critical Success Index” (CSI),

$$\textbf{CSI} = \frac{hits}{hits + misses + false\ alarms}. \tag{6.4}$$

It illustrates the fraction of correct forecasted events to all events. It ranges between 0 and 1 with a perfect value of 1. The CSI is sensitive towards ”hits” but punishes on the other side ”misses” and ”false alarms” simultaneously.

6.1 Neighbourhood Verification

The so-called ”Neighbourhood Verification” approach as described in e.g. EBERT (2006) or CASATI et al. (2004) compares forecast and observed objects deterministically. But in contrast to the classical box-to-box method, before starting the matching, the forecast fields are blurred by using a certain neighbourhood around the specific grid point to be compared (see figure 12). Hence, the observation is matched to the blurred forecast allowing a certain spatial distance between observed and forecasted thunderstorm objects. An object that is correctly forecasted in time but spatially somewhat displaced to the observation would then still be rated as a true prognosis and not as a false one as in the classical box-to-box comparison. The ”Neighbourhood Verification” approach is a sensible concession to thunderstorm forecasting algorithms because thunderstorm cells are small-scale and short-living weather phenomena and therefore difficult to predict precisely in location and time. Typically in the ”Neighbourhood Verification” approach,

the objects are compared for incrementally larger neighbourhoods so that one can determine the scale at which a desired level of skill is attained by the forecast (GILLELAND et al., 2009).

In the present verification we apply the technique of "Multi-event contingency tables" (see ATGER, 2001) within the "Neighbourhood Verification". The traditional 2x2 contingency table uses a single definition of an event, typically whether or not the value of the variable exceeds a given intensity threshold. The "Multi-event contingency table" method not only extends the 2x2 contingency table to several intensity thresholds, it further allows additional dimensions like thresholds on spatial and temporal closeness to be included (EBERT, 2008). Here we use two different dimensions to verify the Cb-LIKE forecasts. For dimension 1 we set several thresholds to determine the forecast objects. These could be indicator or dBZ values. Dimension 2 includes different spatial distances between forecast and observation object which are determined by different neighbourhood sizes. Hence, we evaluate the forecast fields on different scales and find the best setting of threshold and neighbourhood size for Cb-LIKE.

6.2 Setting of the verification

The verification scores are calculated for nine different neighbourhood sizes comprising 1x1 up to 31x31 grid points, see table 5. We call the collective of all neighbourhoods a "neighbourhood ensemble". The thresholds which define the Cb-LIKE forecast objects range from indicator 20 to indicator 80; the thresholds for the synthetic radar field range from 10 to 60 dBZ, see table 6. The threshold value of 37 dBZ is used for a direct comparison with the observed precipitation depicted as Rad-TRAM objects. Again, we name the collective of all thresholds "threshold ensembles".

The verification domain covers the grid points 100-280 in x-direction and 90-390 in y-direction of the COSMO-DE model domain and therefore depicts an area of nearly 400.000 km^2 (grid size is 2.8 km) of Germany and little parts of the neighbouring countries. For this domain, radar data are readily available from DWD.

The time period of the verification encompasses the summer of 2012 starting at 01 June up to 30 September (122 days). Eight model runs per day under application of one up to six hours of forecast imply overall 5.856 calculation steps. The verification results are averaged over all calculation steps. For the object comparison we applied the "Multi-event contingency tables" approach featuring a minimum overlap criterion (see ATGER, 2001). Hence, only one grid point of overlap between observation and forecast plus neighbourhood is required for a positive object matching. Only forecast objects featuring at least 20 grid points, a minimum area of 156,8 km^2 are taken into consideration. Hence, the smallest forecast object is about the same size as the smallest possible Rad-TRAM heavy precipitation cell which consists of at least 21 pixels with a resolution of 2x2 km^2 and cover therefore a surface of 84 km^2, cf. Chapter 2.

6.3 Results

Figures 13 (a) and (b) show the calculated BIAS for Cb-LIKE and the synthetic radar field. Both figures illustrate the results for all 63 possible pairs of the two ensembles. For indicators of 20 and 30 the BIAS attains values > 1 for the most neighbourhoods. This means that the Cb-LIKE algorithm tends to overforecasting for these thresholds when applied for defining the forecast objects. For indicators of 40 or higher the BIAS attains consistently values < 1. The higher the thunderstorm indicator the lower is the total number of forecasted thunderstorm objects. The alteration of the BIAS dependent on the different neighbourhood sizes shows also some interesting characteristics. It becomes smaller for larger edge lengths of the neighbourhoods referred to the indicators of 20 and 30. A possible explanation for this behaviour is a merging of many small into few large forecast objects due to the growing neighbourhoods. This causes the total number of Cb-LIKE forecast objects and therefore the BIAS to decrease. In contrast the BIAS attains higher values for larger neighbourhoods associated with indicators of 40 or higher. In this case merging is more unlikely because Cb-LIKE objects are smaller on average

and exhibits a larger spatial distance among themselves due to the higher thresholds. A probable explication is an increasing number of Cb-LIKE objects which exceeds the minimum threshold of 20 grid points because of the increasing neighbourhood sizes before being sorted out. As a consequence, these objects are now additionally taken into account for the verification and cause the BIAS to rise. A BIAS = 1 can obviously be reached with an indicator threshold between 30 and 40.

The synthetic radar field features a similar distribution of the BIAS. The thresholds from 10 up to 40 dBZ reveal distinct more forecast than observed objects (BIAS > 1). This is valid over the complete neighbourhood ensemble associated to the low dBZ thresholds. The BIAS values are also noticeably higher than for the Cb-LIKE output. For dBZ values of 50 or higher the BIAS drops below 1 due to a rapid decrease of the total number of forecasted synthetic radar objects of that high intensity. With respect to the different neighbourhood sizes the alteration of the BIAS for the synthetic radar field follows the same rules as previously explained for Cb-LIKE. For the 37-dBZ threshold, the one taken in Rad-TRAM, the COSMO-DE model calculates 1.69 to 1.19 time more heavy precipitation cells than observed in the complete neighbourhood ensemble. A perfect BIAS probably lies close to the threshold of 40 dBZ.

In order to assess the quality of forecasts from Cb-LIKE and the synthetic radar field, one should search for the thresholds which result in a BIAS value of 1. Then, both prognosis fields calculate an approximate equal number of objects and can be compared with regard to their general forecasting performance. As illustrated in figure 14, the required thresholds for Cb-LIKE and the synthetic radar field are an indicator of 30 and 41 dBZ, respectively. Averaged over the neighbourhood ensemble the BIAS attains a value of approximately 1 for both data fields. Figure 15 illustrates the results for the FAR for these two thresholds. It can be seen that the Cb-LIKE algorithm exhibits a lower ”False Alarm Ratio” (less false alarms) than the synthetic radar field over all neighbourhoods. Figure 16 presents the final results for the ”Probability of

Detection". Cb-LIKE shows higher POD values over the complete neighbourhood ensemble than the COSMO-DE synthetic radar field. Hence the observed Rad-TRAM heavy precipitation cells are better forecasted with Cb-LIKE. Figure 17, finally, exhibits the "Critical Success Index" and illustrates that the fraction of correctly forecasted events to all events is higher for Cb-LIKE than for the synthetic radar field. In summary, the comparison of both data fields for an approximately equal BIAS value of 1 illustrates a better performance of Cb-LIKE. The new forecast algorithm features less false alarms, more correctly predicted Rad-TRAM heavy precipitation cells and also a higher portion of correct forecasted events. Nevertheless, in all three verification scores it is observed that the quality of forecasts based on the synthetic radar field catches up with the Cb-LIKE forecasts with increasing neighbourhood sizes. The larger the neighbourhood the less the spatial error between forecast and observation influences the verification results.

6.4 Thunderstorm probability and neighourhood size

Eventually, the comprehensive results of the verification enable the transformation of the Cb-LIKE thunderstorm indicators into thunderstorm probabilities. We apply the calculated FAR for this purpose. This verification score describes the average probability of a forecasting object being a false alarm. Thereby (1-FAR) represents the mean probability of a Cb-LIKE forecast object hitting a Rad-TRAM observation. If, for example, the Cb-LIKE objects defined with an indicator threshold of 50 feature a FAR of 0,7 (70 %), 30 percent of these objects would hit an observation. Hence, the indicator of 50 is equivalent to a mean thunderstorm probability of 30 %.

Table 7 (left table) presents the mean "False Alarm Ratios" for all indicator thresholds (20 up to 80) for the smallest neighbourhood of the neighbourhood ensemble. The average thunderstorm probability in percent (right column of the table) is calculated by $(1 - FAR) * 100$. The table illustrates an increase of the thunderstorm probability for higher indicators. It is only 25 % for

an indicator threshold of 20 but rises up to 51 % for an indicator of 50. The growth of the thunderstorm probability for higher indicators is a sensible characteristic due to the fact that higher indicators suggest a state of the atmosphere which is more prone to the development of thunderstorms.

Table 7 (right table) shows the same quantities for a neighbourhood with an edge length of 53.2 km. As one might expect, the thunderstorm probabilities reach higher values compared to the smaller neighbourhood size. Indicator thresholds of 20/50/80 represent mean thunderstorm probabilities of 53/79/90 %, respectively. The higher values follow from a higher fuzziness of the Cb-LIKE forecasts caused by the larger neighbourhood. In the case at hand the Cb-LIKE forecasts can be located within a distance of 53.2 km to the observation and still be rated as a matching forecast. The neighbourhood size should always be chosen according to the application in mind. For example, the size of about 50 km is an appropriate value when considering airport environments, the so-called "Terminal Maneuvering Area".

7 Conclusions and outlook

Cb-LIKE calculates probabilistic thunderstorm forecasts up to six hours and follows up the deterministic nowcasting for time horizons between 0 and 1 hour. The two prognosis regimes are seamlessly coupled by a "Best-Member-Selection". This tool chain fulfils the requirements as pointed out by aviation stakeholders to provide observations and forecasts of thunderstorm cell developments from now to several hours ahead in time. We chose the COSMO-DE model which is operated by the German Meteorological Service (see BALDAUF et al., 2011) as data source. This nonhydrostatic numerical weather model provides one of the highest grid resolution among all available operational models in Middle Europe and in addition a full calculation of deep convection (no parametrization). Hence it is the first option for calculating probabilistic thunderstorm forecasts in Middle Europe. The first section of the Cb-LIKE algorithm includes

the newly developed "Best-Member-Selection". This part allows the automatic choice of the specific COSMO-DE model run out of a (time-lagged) ensemble which matches best the current weather situation. We do the matching by a comparison between observed (Rad-TRAM) and synthetic (COSMO-DE model) radar objects. Afterwards the model output fields of the selected model run are combined within a newly developed fuzzy logic system. Fuzzy logic enables generally the conjunction of several data fields applying human reasoning and expert knowledge simultaneously. Altogether we utilize four COSMO-DE model output fields (CAPE, OMEGA at 500 hPa, Cloud top temperature and Radar reflectivity) within the fuzzy logic system. A so-called thunderstorm indicator for each grid point of the COSMO-DE model domain and for all output times up to 6 hours in one hour intervals represents the final output. This quantity can attain values between $11.\overline{6}$ and $88.\overline{3}$. The higher the value the more the input model parameters indicate the development/existence of thunderstorms.

A thorough verification over the summer period of 2012 shows the general quality of the Cb-LIKE forecasts. For this purpose we verified Cb-LIKE forecasts and the synthetic radar field of the COSMO-DE model against heavy precipitation cells observed by Rad-TRAM by using a deterministic object comparison under application of a so-called "Neighbourhood Verification" approach (see e.g. EBERT, 2008) and "Multi-event contingency tables" (see ATGER, 2001). The results show that the new algorithm Cb-LIKE is an improvement in forecasting thunderstorms on the basis of the COSMO-DE model. Eventually, the verification results allow to map the Cb-LIKE thunderstorm indicator values into thunderstorm probability values.

Besides the first setting of the Cb-LIKE algorithm presented in this publication several enhancements of the system are possible for the future:

- Application of the COSMO-DE EPS instead of the COSMO-DE model as data source (COSMO-DE EPS = 20 members per run; THEIS et al., 2005) or KENDA ("Km-Scale Ensemble-Based Data Assimilation"; REICH et al. (2011) - future project for the improve-

ment of the forecasting of convective situations).

- Utilization of more meteorological parameters within the fuzzy logic system which are also indicators of a possible thunderstorm development: e.g. vertical wind shear, solar radiation (magnitude), orographic forcing or jetstream divergence (vertical wind, stability of atmosphere).

- Experiments with different shapes of the fuzzy input sets and various sizes of the transition areas: tuning of the fuzzy logic system could lead to an improvement of the Cb-LIKE forecast quality.

In conclusion, Cb-LIKE with its innovative fuzzy logic approach shows promising results in providing probabilistic thunderstorm forecasts up to 6 hours. The comparison with the synthetic radar data shows that this new algorithm is a sensible method for thunderstorm forecasting on the basis of COSMO-DE model data. For an easy and appealing visualisation, the Cb-LIKE forecasts are implemented in the WxFUSION system (TAFFERNER et al., 2008; FORSTER and TAFFERNER, 2008) with its graphical user interface for visualisation and allows a straightforward comparison of the new forecasts with other data sources for scientific purposes.

Acknowledgments

This work was done at the Department of Traffic Meteorology of the Institute for Atmospheric Physics within the framework of the DLR-Project WOLV. First I would like to express my special gratitude to Thomas Gerz and Arnold Tafferner for their contribution to the development of Cb-LIKE. In addition I have to thank my colleagues at the institute for their help to realize this project. Finally I want to thank the DWD for providing access to the COSMO-DE data bases.

References

ATGER, F., 2001: Verification of intense precipitation forecasts from single models and ensemble prediction systems. – Nonlinear Processes in Geophysics **8**, 401–417.

BALDAUF, M., J. FÖRSTER, S. KLINK, T. REINHARDT, C. SCHRAFF, A. SEIFERT, K. STEPHAN, 2011: Kurze Beschreibung des Lokal-Modells Kürzestfrist COSMO-DE (LMK) und seiner Datenbanken auf dem Datenserver des DWD. – Stand 31.03.2011, Deutscher Wetterdienst, Geschäftsbereich Forschung und Entwicklung, Postfach 100465, D-63004 Offenbach

BETZ, H. D., K. SCHMIDT, W. P. OETTINGER, B. MONTAG, 2008: Cell-Tracking with Lightning Data from LINET. – Adv. Geosci **17**, 55–61.

CASATI, B., G. ROSS, D. B. STEPHENSON, 2004: A new intensity-scale approach for the verification of spatial precipitation forecasts. – Meteor. Appl. **11**, 141–154.

DIXON, M., G. WIENER, 1993: TITAN: Thunderstorm Identification, Tracking, Analysis, and Nowcasting - A Radar-based Methodology. – J. Atmos. Oceanic Technol. **10**(6), 785–797.

DJAM, X. Y., G. M. WAJIGA, Y. H. KIMBI, N. V. BLAMAH, 2011: A Fuzzy Expert System for the Management of Malaria. – Int. J. Pure Appl. Sci. Technol. **5(2)**, 84–108.

DOSWELL III, C. A., R. DAVIES-JONES, D. L. KELLER, 1990: On Summary Measures of Skill in Rare Event Forecasting Based on Contingency Tables. – Wea. Forecasting **5**, 576–585.

EBERT, E. E., 2006: Fuzzy Forecast Verification. – MAP D-PHASE workshop, 6-8 Novemer 2006.

EBERT, E. E., 2008: Fuzzy Verification of high-resolution gridded forecasts: a review and proposed framework. – Meteorol. Appl. **15**, 51–64.

EILTS, M. D., COAUTHORS, 1996: Severe weather warning decision support system. – Preprints, 18th Conf. on Severe Local Storms, San Francisco, CA, Amer. Meteor. Soc. 536–540.

EUROCONTROL, 2011: Performance Review Commission, 2011: Performance Review Report 2011. – 128pp, verfügbar unter https://www.eurocontrol.int/sites/default/files/content/documents/single-sky/pru/publications/prr/prr-2011.pdf

FORSTER, C., A. TAFFERNER, 2008: An Integrated User-oriented Weather Forecast System for Air Traffic Using Real-Time Observations and Model Data. – ICAS 2008 Proceedings Volume (863), Seiten 1-10. Optimage Ltd., Edinburgh, UK. 26TH CONGRESS OF THE INTERNATIONAL COUNCIL OF THE AERONAUTICAL SCIENCES, 2008-09-14, Anchorage, AK (USA). ISBN 0-9533991-9-2.

GILLELAND, E., D. AHIJEVYCH, B. G. BROWN, B. CASATI, E. E. EBERT, 2009: Intercomparison of Spatial Forecast Verification Methods. – Wea. Forecasting **24**, 1416–1430.

HANSEN, B., 2007: A Fuzzy Logic-Based Analog Forecasting System for Ceiling and Visibility. – Wea. Forecasting **22**, 1319–1330.

JANTZEN, J., 1998: Tutorial on fuzzy logic. Technical report, Technical University of Denmark, Oersted-DTU, Automation, Bldg 326, 2800 Kongens Lyngby, DENMARK. Tech. Report no 98-E 868 (logic), revised 22 March 2006.

KEIS, F., 2015: WHITE - Winter hazards in terminal environment: An automated nowcasting system for Munich Airport. – Meteorol. Zeitschrift **24**, 61–82.

KLIR, G., B. YUAN, 1995: Fuzzy Sets and Fuzzy Logic: Theory and Applications – Prentice Hall New Jersey.

652 KOBER, K., A. TAFFERNER, 2009: Tracking and Nowcasting of convective cells using remote
653 sensing data from radar and satellite. – Meteorol. Zeitschrift **1**, 75–84.

654 KOBER, K., G. C. CRAIG, B. KEIL, A. DÖRNBRACK, 2012: Blending a probabilistic now-
655 casting method with a high-resolution numerical weather prediction ensemble for convective
656 precipitation forecasts. – Q. J. R. Meteorol. Soc. **138**, 755–768.

657 LEIGHTON, Q., 2006: Modeling and Simulation Needs for Next Generation Air Transportation
658 System Research. – AIAA Modeling and Simulation Technologies Conference and Exhibit,
659 Keystone, CO, 21-24 August 2006, AIAA 2006-6109, 1-8.

660 MANZATO, A., 2005: The Use of Sounding-Derived Indices for a Neural Network Short-Term
661 Thunderstorm Forecast. – Wea. Forecasting **20**, 896–917.

662 MUELLER, C., T. SAXEN, R. ROBERTS, J. , T. BETANCOURT, S. DETTLING, N. OIEN, J. YEE,
663 2003: NCAR Auto-Nowcast System. – Wea. Forecasting **18**, 545–561.

664 MUELLER, C. K., J. W. WILSON, N. A. CROOK, 1993: The Utility of Sounding and Mesonet
665 Data to Nowcast Thunderstorm Initiation. – Wea. Forecasting **8**, 132–146.

666 MURTHA, J., 1995: Applications of Fuzzy Logic In Operational Meteorology. – Scientific
667 Services and Professional Developement Newsletter, Canadian Forces Weather Service 42–
668 54.

669 REICH, H., A. RHODIN, C. SCHRAFF, 2011: LETKF for the nonhydrostatic regional
670 model COSMO-DE Newsletter11, Consortium for Small-scale Modelling. http://www.cosmo-
671 model.org/content/model/documentation/newsLetters(accessed 13 February 2014).. –

672 ROSS, T. J., 2010: Fuzzy logic With Engineering Applications – Chichester, U.K. John Wiley
673 & Sons Ltd, 3rd ed.

SCHÄTTLER, U., G. DOMS, C. SCHRAFF, 2013: A Description of the Nonhydrostatic Regional COSMO-Model. – Printed at Deutscher Wetterdienst, P.O. Box 100465, 63004 Offenbach, Germany.

SCHMEITS, M. J., K. J. KOK, D. H. P. VOGELEZANG, 2005: Probabilistic Forecasting of (Severe) Thunderstorms in the Netherlands Using Model Output Statistics. – Wea. Forecasting **20**, 134–148.

SCHULZ, J. P., U. SCHÄTTLER, 2010: Kurze Beschreibung des Lokal-Modells Europa COSMO-EU (LME) und seiner Datenbanken auf dem Datenserver des DWD. – Stand 20.05.2010, Deutscher Wetterdienst, Geschäftsbereich Forschung und Entwicklung, Postfach 100465, D-63004 Offenbach.

TAFFERNER, A., H. MANNSTEIN, T. PACCAGNELLA, C. MARSIGLI, A. MONTANI, F. NEROZZI, 2002: In: 10th Conference on Mountain Meteorology, Seiten 201-204. 10th Conference on Mountain Meteorology, Park City, UT, USA, 17-21 June 2002.. –

TAFFERNER, A., M. HAGEN, C. KEIL, T. ZINNER, H. VOLKERT, C. FORSTER, 2008: Development and Propagation of Severe Thunderstorms in the Upper Danube Catchment Area: Towards an Integrated Nowcasting and Forecasting System using Real-time Data and High-resolution Simulations. – Meteorology and Atmospheric Physics **101**, 211–227.

THEIS, S. E., A. HENSE, U. DAMRATH, 2005: Probabilistic precipitation forecasts from a deterministic model: a pragmatic approach. – Meteorol. Appl. **12**, 257–268.

TREPTE, S., 2011: WarnMOS - A MOS-based weather warning system. – 11th EMS Annual Meeting/10th European Conference on Applications of Meteorology (ECAM), Berlin, Germany, 12 - 16 September

UMOH, U. A., E. O. NWACHUKWU, O. E. OKURE, 2010: Fuzzy Rule-based Framework for Effective Control of Profitability in a Paper Recycling Plant. – Global Journal of Computer Science and Technology **10**(10), 56–67.

WEIGL, E., S. KLINK, O. KOHLER, T. REICH, W. ROSENOW, P. LANG, C. PODLASLY, T. WINTERRATH, D. MAJEWSKI, J. LANG, 2005: Abschlussbericht Projekt RAD-VOROP: RadargestÃ¼tzte, zeitnahe Niederschlagsvorhersage fÃ¼r den operationellen Einsatz (Niederschlag-Nowcasting-System). – Technical report **Deutscher Wetterdienst, Abteilung Hydrometeorologie.**

WILKS, D. S., 2006: Statistical Methods in the Atmospheric Sciences – International geophysics series, San Diego, New York: Academic Press, Elsevier, 2nd ed.

WONG, M., W. WONG, E. LAI, 2006: From SWIRLS to RAPDIS: Nowcasting applications developement in Hong Kong.. – Proc. PWS Workshop on Warnings of Real-Time Hazards by Using Nowcasting Technology, Sidney, Australia, WMO, 9-13.

ZINNER, T., H. MANNSTEIN, A. TAFFERNER, 2008: Cb-Tram: Tracking and monitoring severe convection from onset over rapid development to mature phase using multi-channel Meteosat-8 SEVIRI data. – Meteorol. Atmos. Phys. **101**, 191–210.

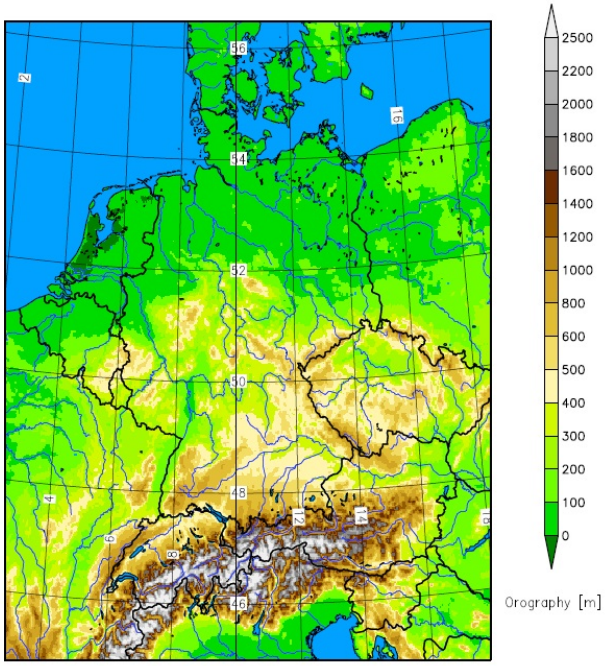


Figure 1: The whole domain covered by the COSMO-DE model which is driven by German Meteorological Service (DWD) (Source: BALDAUF et al., 2011).

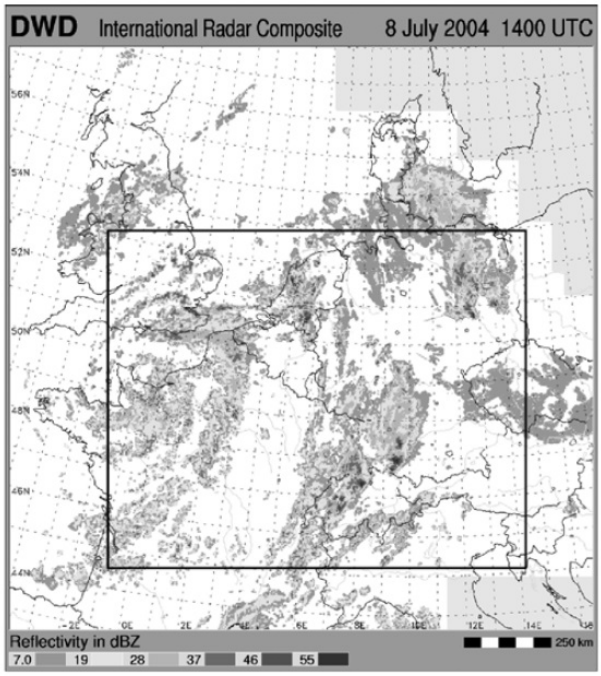


Figure 2: European radar composite at 08 June 2004 at 1400 UTC. The black framed square illustrates the domain covered by Rad-TRAM (Source: KOBER and TAFFERNER, 2009).

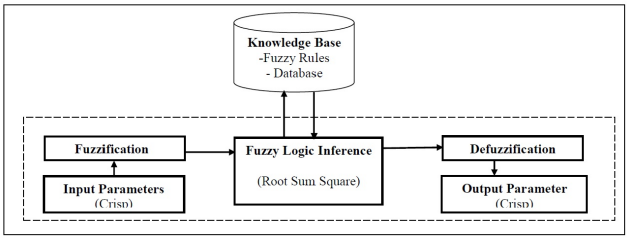


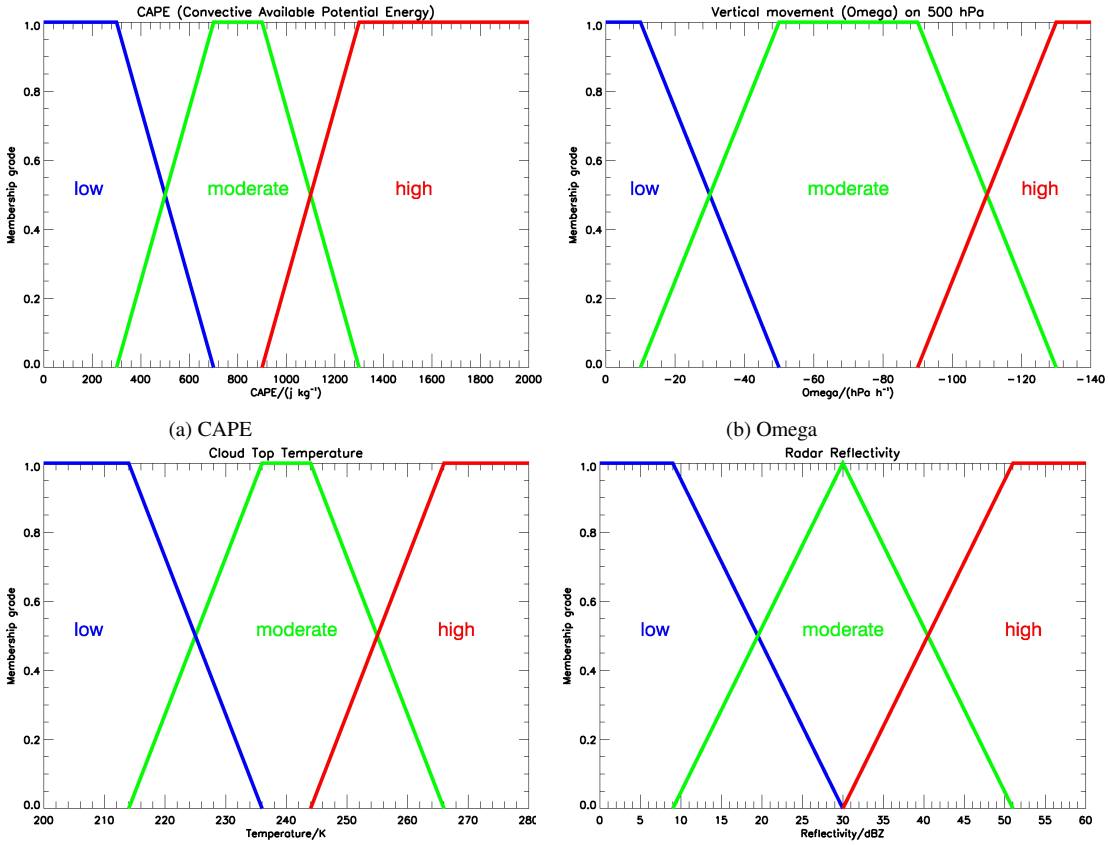
Figure 3: Outline of a general fuzzy logic system which is divided into three main working steps: fuzzification, fuzzy logic inference and defuzzification (Source: DJAM et al., 2011).

	low	moderate	high
CAPE/[j/kg]	0 to 700	300 to 1300	900 to 2000
Omega/[hPa/h]	0 to -50	-10 to -130	-90 to -140
Radar/dBZ	0 to 30	9 to 51	30 to 60
CloudTop/K	200 to 236	214 to 266	244 to 280

Table 1: Value ranges of the four input parameters which are covered by the input sets low, moderate and high, respectively.

Input sets	**Virtual score**
CAPE$_{high}$, Omega$_{high}$, Radar$_{high}$, CloudTop$_{low}$	+1
CAPE$_{moderate}$, Omega$_{moderate}$, Radar$_{moderate}$, Satellit$_{moderate}$	0
CAPE$_{low}$, Omega$_{low}$, Radar$_{low}$, Satellit$_{high}$	-1

Table 2: Outline of the assignment of the virtual scores (x_i) (-1/0/1) to the fuzzy input sets.



(a) CAPE

(b) Omega

(c) Cloud top temperature

(d) Radar reflectivity

Figure 4: Three membership functions (low, moderate, high) assign the four crisp input parameters CAPE (a), Omega (b), cloud top temperature (c), and radar reflectivity (d) membership grades between 0 and 1 to the associated fuzzy input set(s), respectively. On each figure the abscissa comprises a certain value range whereas the ordinate includes the membership grade from 0 to 1.

Lower limit	Conjunction with output sets	Upper limit
-1.0 <= m	Very low	m < -0.6
-0.6 <= m	low	m < -0.2
-0.2 <= m	moderate	m <= 0.2
0.2 < m	high	m <= 0.6
0.6 < m	very high	m <= 1.0

Table 3: Outline of the conjunction of a decision rule to a specific output set dependent on the weighted average m.

Output set	Number of rules
Very low	5
Low	26
Moderate	19
High	26
Very high	5

Table 4: Exact conjunction of all 81 "if...then" decision rules to the fuzzy output sets within Cb-LIKE.

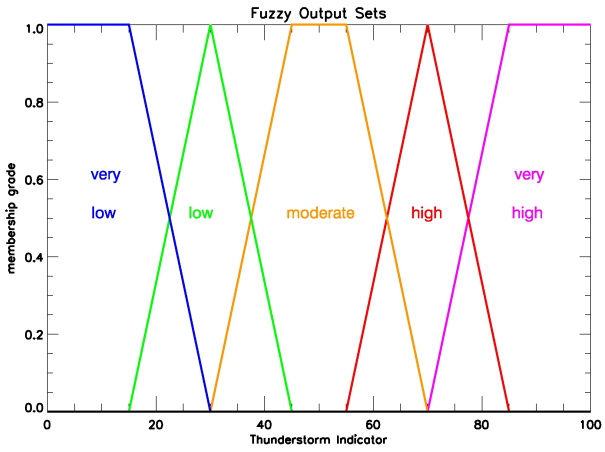
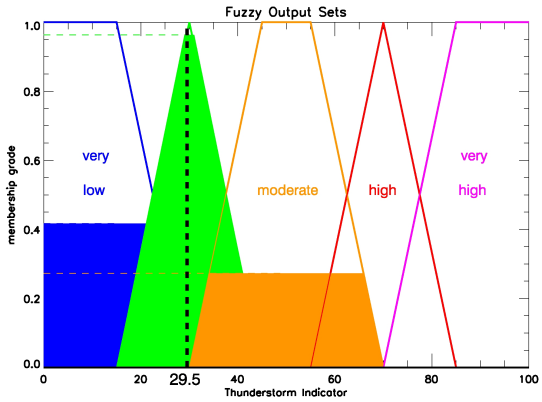
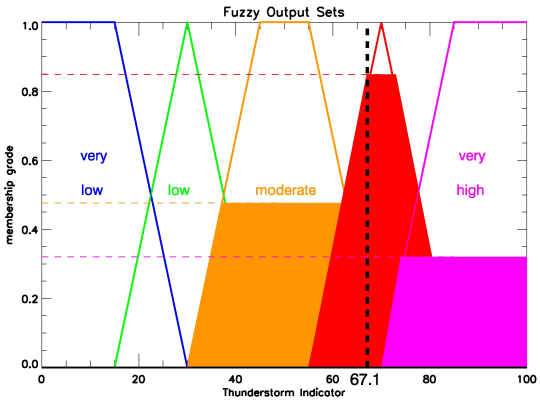


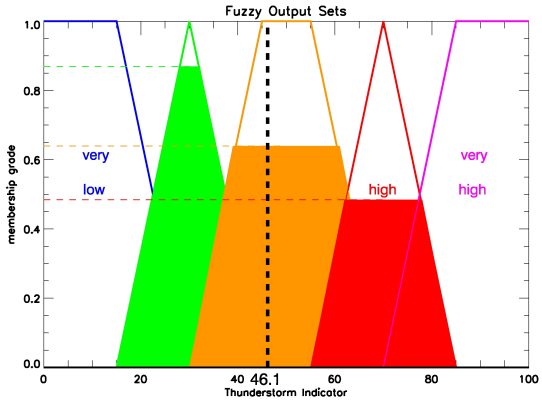
Figure 5: Five fuzzy output sets (very low, low, moderate, high, very high) cover a range from 0 to 100 on the abscissa for the calculation of the final output of Cb-LIKE: the thunderstorm indicator. The ordinate comprises the membership grade from 0 up to 1.



(a) Weighted input sets lead to a low thunderstorm indicator of 29.5

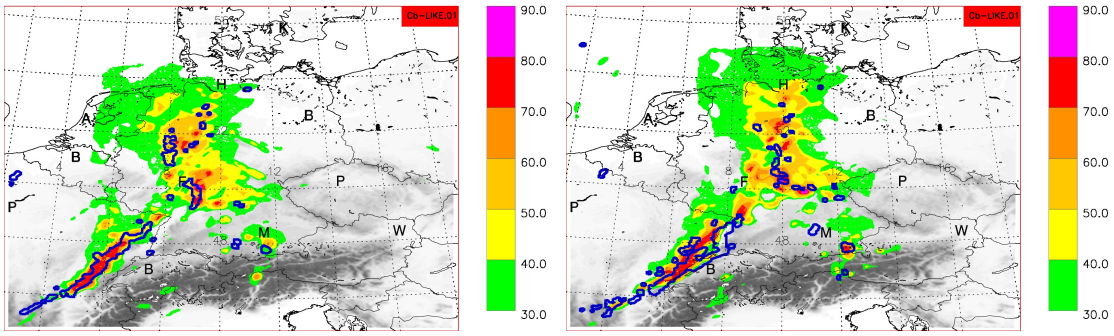


(b) Weighted input sets lead to a high thunderstorm indicator of 67.1



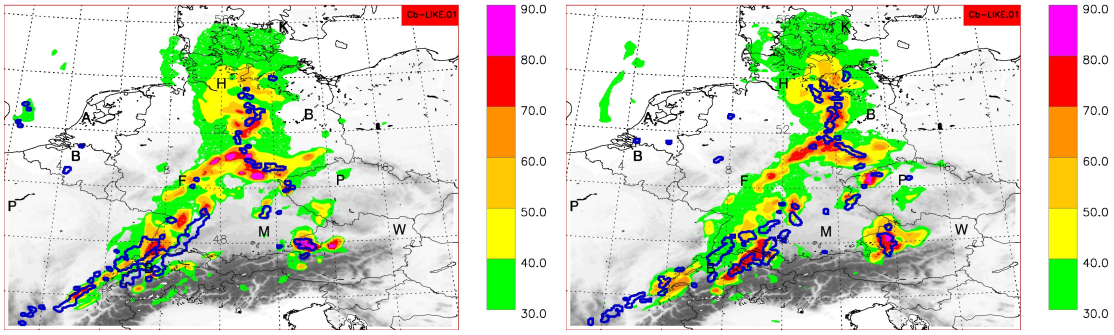
(c) Weighted input sets lead to a moderate thunderstorm indicator of 46.1

Figure 6: The diagrams (a) to (c) show the different thunderstorm indicators as obtained from the five fuzzy output sets featuring certain membership grades caused by the different input parameters. The membership grades are illustrated as coloured surfaces. In (a) the input parameters CAPE = 450.0 j/kg, Omega = -45.0 hPa/h, Radar reflectivity = 23.0 dBZ and Cloud top temperature = 260.0 K lead to a rather low indicator of 29.5. In (b) the high indicator of 67.1 is calculated using the following input parameters: CAPE = 950.0 j/kg, Omega = -98.0 hPa/h, Radar reflectivity = 41.0 dBZ and Cloud top temperature = 228.0 K. In figure (c) a moderate indicator of 46.1 bases on the input parameters CAPE = 450.0 j/kg, Omega = -98.0 hPa/h, Radar reflectivity = 41.0 dBZ and Cloud top temperature = 260.0 K.



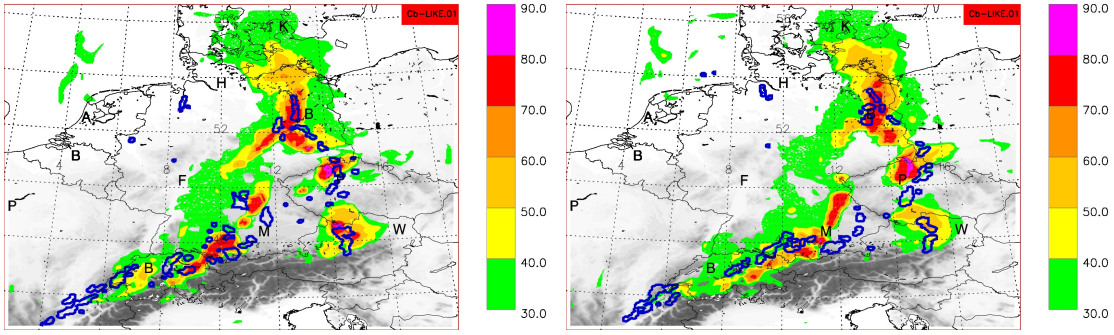
(a) Storm activity at 1300 UTC plus corresponding Cb-LIKE forecast from 1200 UTC

(b) Storm activity at 1400 UTC plus corresponding Cb-LIKE forecast from 1200 UTC



(c) Storm activity at 1500 UTC plus corresponding Cb-LIKE forecast from 1200 UTC

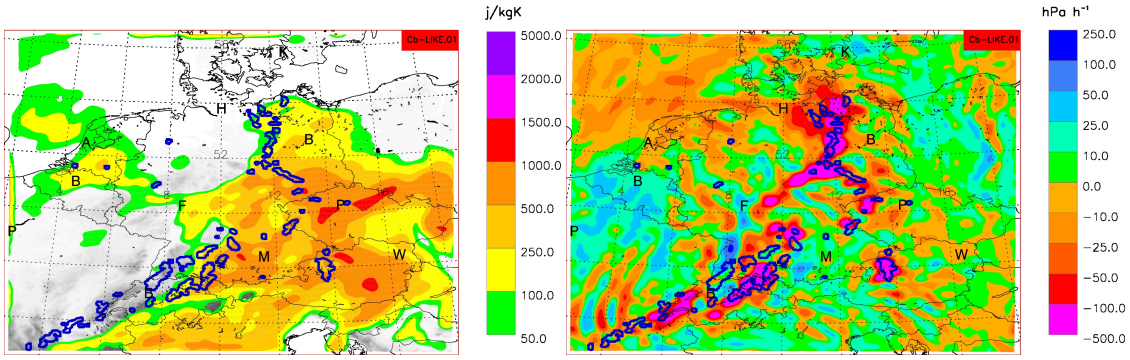
(d) Storm activity at 1600 UTC plus corresponding Cb-LIKE forecast from 1200 UTC



(e) Storm activity at 1700 UTC plus corresponding Cb-LIKE forecast from 1200 UTC

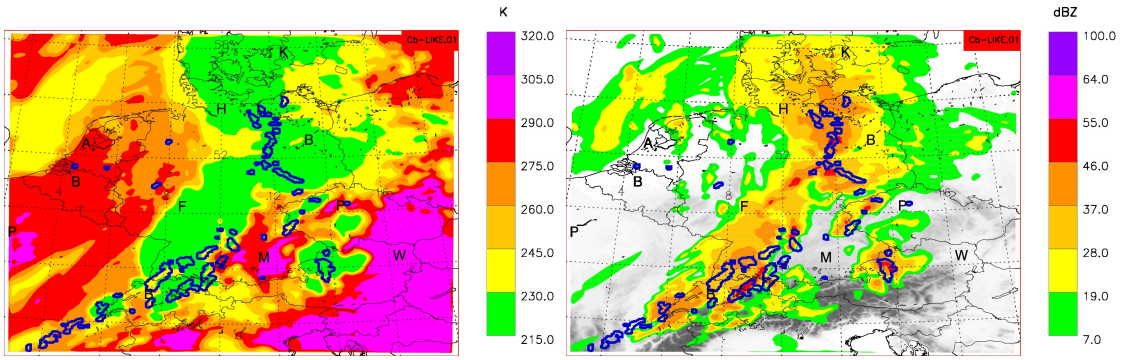
(f) Storm activity at 1800 UTC plus corresponding Cb-LIKE forecast from 1200 UTC

Figure 7: The diagrams (a) to (f) show the observed thunderstorm activity (Rad-TRAM data) from 1300 up to 1800 UTC and the corresponding one to six hours Cb-LIKE indicator prognoses from 1200 UTC on 22 June 2011. The thunderstorm activity is represented by blue contour lines whereas the Cb-LIKE forecasts are displayed as coloured areas based upon the COSMO-DE model run from 1200 UTC. The figures encompass the whole domain of the COSMO-DE model.



(a) CAPE

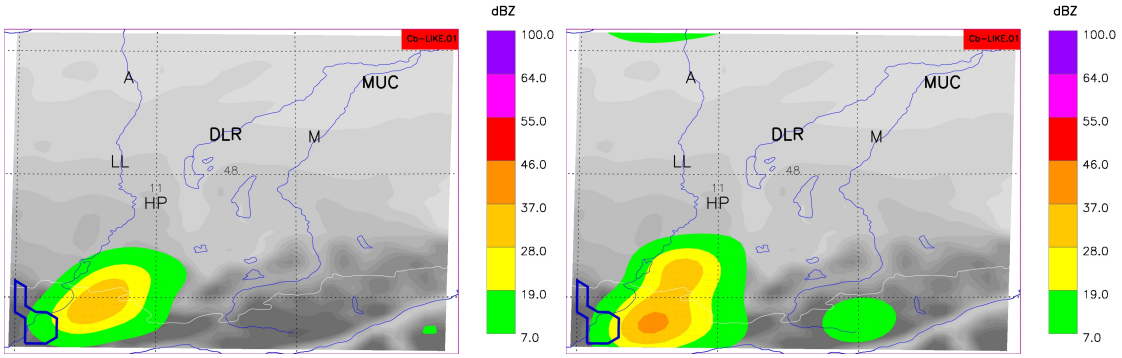
(b) OMEGA on 500 hPa level



(c) Cloud top temperature

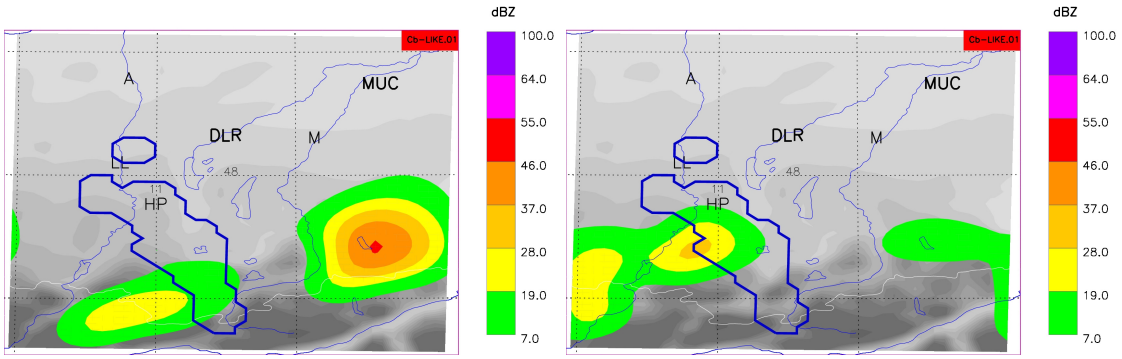
(d) Synthetic radar reflectivity

Figure 8: The diagrams (a) to (d) show the observed thunderstorm activity (Rad-TRAM data) at 1600 UTC on 22 June 2011 plus the corresponding COSMO-DE parameters (CAPE, OMEGA (500 hPa), Cloud top temperature and Radar reflectivity) which are used within the Cb-LIKE algorithm. The parameters are extracted from the 1200 UTC model run.



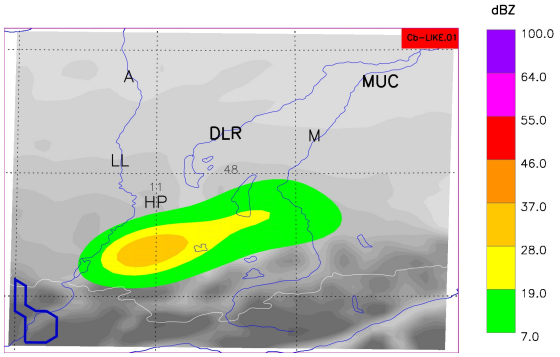
(a) Storm activity at 1400 UTC plus corresponding synthetic radar reflectivity field based on COSMO-DE model run from 1200 UTC

(b) As (a) for 0900 UTC COSMO-DE model run



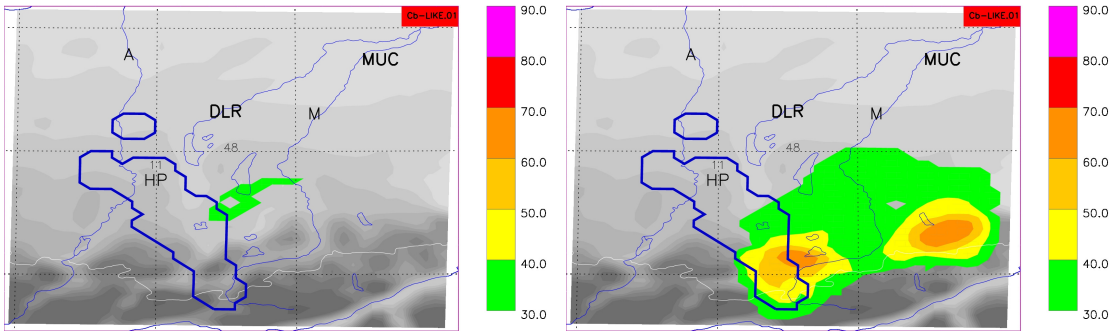
(c) As (a) for 0600 UTC COSMO-DE model run

(d) As (a) for 0300 UTC COSMO-DE model run



(e) As (a) for 0000 UTC COSMO-DE model run

Figure 9: The diagrams (a) to (e) show the observed thunderstorm activity (Rad-TRAM data) and the corresponding synthetic radar reflectivity fields of the last five COSMO-DE model runs (1200 - 0000 UTC) of the time-lagged ensemble in upper Bavaria at 1500 UTC on 05 August 2011. The thunderstorm activity is represented by blue contour lines whereas the reflectivity values are displayed as coloured areas.



(a) Storm activity at 1500 UTC plus corresponding Cb-LIKE forecast based on the COSMO-DE model run from 1200 UTC

(b) As (a) for 0900 UTC COSMO-DE model run

(c) As (a) for 0600 UTC COSMO-DE model run

(d) As (a) for 0300 UTC COSMO-DE model run

(e) As (a) for 0000 UTC COSMO-DE model run

Figure 10: The figures (a) to (e) show the observed thunderstorm activity (Rad-TRAM data) at 1500 UTC on 05 August 2011 and the corresponding Cb indicator prognoses based on the last five COSMO-DE model runs (1200 - 0000 UTC) of the time-lagged ensemble. The thunderstorm activity is represented by blue contour lines whereas the Cb indicator prognoses are displayed as coloured areas. The figures encompass mainly upper Bavaria.

		Observation: Yes	Observation: No
Forecast	Yes	*hits*	*false alarms*
	No	*misses*	*correct negatives*

Figure 11: Four different events are possible within a deterministc object comparison: hits, false alarms, misses and correct negatives. This kind of table represents a so-called 2x2 contingency table (adapted from WILKS, 2006).

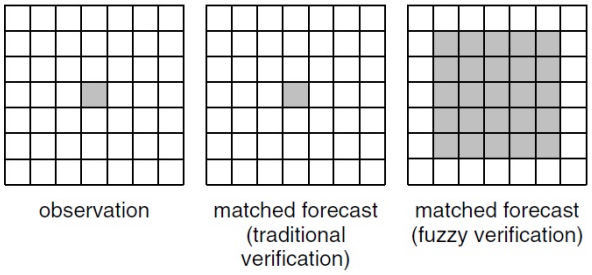


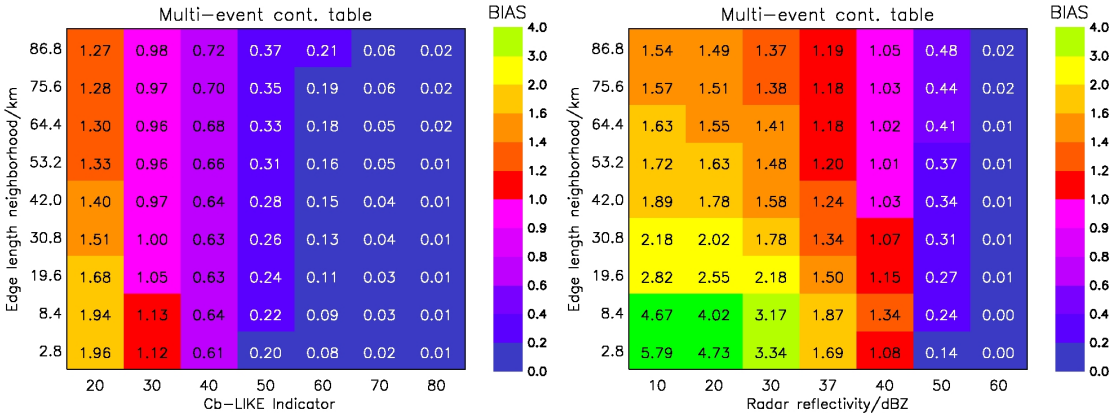
Figure 12: Within the traditional verification a box-to-box comparison between observation and forecast is applied (left/middle figure). The neighbourhood verification approach uses additionally a defined neighbourhood around the box to be evaluated (right figure) within the forecast field. As a result the spatial distance between forecast and observation can be taken into account within the verification. Source of the figure: EBERT (2008).

Size in grid points	Edge length/km
1x1	2.8
3x3	8.4
7x7	19.6
11x11	30.8
15x15	42.0
19x19	53.2
23x23	64.4
27x27	75.6
31x31	86.8

Table 5: Outline of the applied neighbourhood ensemble within the neighbourhood verification. In the left column the nine different neighbourhood sizes in grid points are displayed, the right column contains the corresponding edge lengths in kilometres.

Cb-LIKE indicator	Radar reflectivity/dBZ
20	10
30	20
40	30
50	37
60	40
70	50
80	60

Table 6: Outline of the applied threshold ensembles for Cb-LIKE indicator and the COSMO-DE synthetic radar field within the neighbourhood verification. Seven different indicator values and dBZ values are used for the object defining out of both prognoses fields, respectively.



(a) BIAS for Cb-LIKE indicator prognoses

(b) BIAS for COSMO-DE synthetic radar prognoses

Figure 13: Calculated verification results of the BIAS for the Cb-LIKE indicator forecasts (a) and COSMO-DE synthetic radar prognoses (b) for all 63 threshold and neighbourhood combinations.

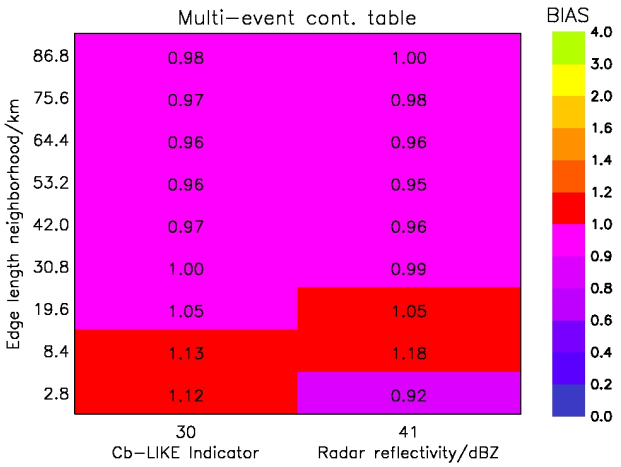


Figure 14: Successful tuning of both data fields to a BIAS = 1. The requested thresholds are 30 for the Cb-LIKE indicator and 41 dBZ for the synthetic radar field. Averaged over all neighbourhood sizes the BIAS attains a value of approximately one for both fields, respectively. The two thresholds were calculated by performing an iterative approach.

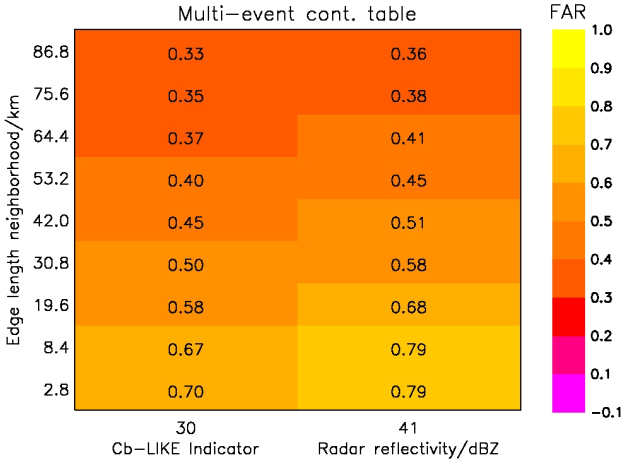


Figure 15: Direct comparison of the FAR between Cb-LIKE and the synthetic radar field of the COSMO-DE model over all neighbourhood sizes. The results of the FAR are calculated under application of the thresholds featuring a mean BIAS = 1 (Indicator = 30, dBZ= 41).

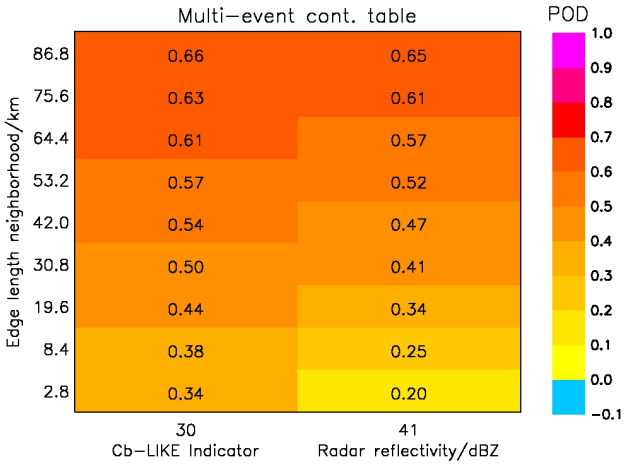


Figure 16: Direct comparison of the POD between Cb-LIKE and the synthetic radar field of the COSMO-DE model over all neighbourhood sizes.

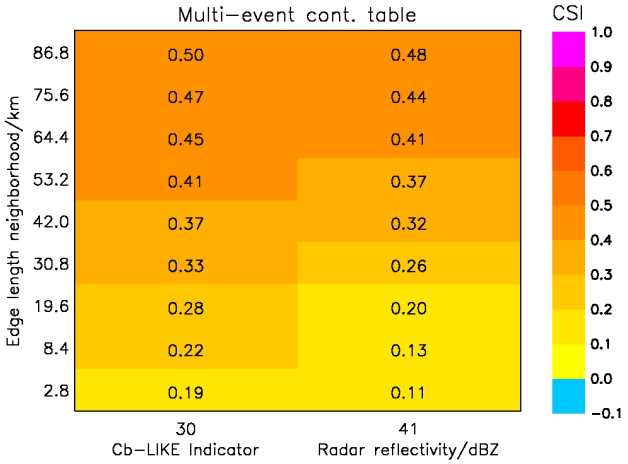


Figure 17: Direct comparison of the CSI between Cb-LIKE and the synthetic radar field of the COSMO-DE model over all neighbourhood sizes.

Indicator	Mean FAR	Probability
20	0.75	25 %
30	0.70	30 %
40	0.64	36 %
50	0.49	51 %
60	0.45	55 %
70	0.36	64 %
80	0.32	68 %

Indicator	Mean FAR	Probability
20	0.47	53 %
30	0.40	60 %
40	0.35	65 %
50	0.21	79 %
60	0.18	82 %
70	0.14	86 %
80	0.10	90 %

Table 7: Both tables show the transformation of the Cb-LIKE indicator values into probabilities using the FAR out of the verification. The left column shows the different Cb-LIKE indicator thresholds from 20 to 80, the middle one the corresponding results of the FAR, respectively. The right column displays the corresponding thunderstorm probabilities which are calculated by $(1 - FAR) * 100$. The edge lengths of the used neighbourhoods are 2.8 km (left table) and 53.2 km (right table).